A prospective evaluation of the predictive value of faecal calprotectin in quiescent Crohn's disease.

Naismith, Graham D.[1];Smith, Lyn A.[1]; Barry, Sarah J.E.[2];Munro, Joanna I.[1];Laird, Susan [1]; Rankin, Karen[3]; Morris, Allan J.[1]; Winter, Jack W.[1]; Gaya, Daniel R.[1]

1 Department of Gastroenterology, Glasgow Royal Infirmary, Glasgow, UK

2 Robertson Centre for Biostatistics, Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK

3 Department of Clinical Biochemistry, Glasgow Royal Infirmary, Glasgow, UK

Corresponding author

Dr Daniel R Gaya | Consultant Gastroenterologist
Gastroenterology Unit | Glasgow Royal Infirmary
G4 0SF
Tel 0141-211-4290 | Fax 0141-211-5162
daniel.gaya@ggc.scot.nhs.uk | daniel.gaya@nhs.net

Running Title – Calprotectin to predict relapse in quiescent Crohn's

Keywords – Crohn's disease, inflammatory bowel disease, diagnostic tests, inflammation, small intestine, large intestine, stomach and duodenum, ROC, Kaplan-Meier, Cox model

ABSTRACT

Background: The faecal calprotectin (FC) test is a non-invasive marker for gastrointestinal inflammation.

Aim: To determine whether higher FC levels in individuals with quiescent Crohn's disease are associated with clinical relapse over the ensuing 12 months.

Methods: A single centre prospective study was undertaken in Crohn's disease patients in clinical remission attending for routine review. The receiver operating characteristic (ROC) curve for the primary endpoint of clinical relapse by 12 months, based on FC at baseline, was calculated. Kaplan-Meier curves of time to relapse were based on the resulting optimal FC cutoff for predicting relapse.

Results: Of 97 patients recruited, 92 were either followed up for 12 months without relapsing, or reached the primary endpoint within that period. Of these, 10 (11%) had relapsed by 12 months. The median FC was lower for non-relapsers, 96µg/g (IQR 39-237), than for relapsers, 414µg/g (IQR 259-590), (p=0.005). The area under the ROC curve to predict relapse using FC was 77.4%. An optimal cutoff FC value of 240µg/g to predict relapse of quiescent Crohn's had sensitivity of 80.0% and specificity of 74.4%. Negative predictive value was 96.8% and positive predictive value was 27.6%. FC≥240µg/g was associated with likelihood of relapse 5.7 (95% CI 1.9-17.3) times higher within 2.3 years than lower values (p=0.002).

Conclusions: In this prospective dataset, FC appears to be a useful, non-invasive tool to help identify quiescent Crohn's disease patients at a low risk of relapse over the ensuing 12 months. FC of 240µg/g was the optimal cutoff in this cohort.

.

BACKGROUND

Calprotectin is a calcium and zinc binding protein found in the cytosol of neutrophils. It is released at times of cell damage in the gastrointestinal (GI) tract and is resistant to enzymatic degradation allowing for measurement in faecal samples. The faecal calprotectin (FC) test has been shown to correlate well with faecal excretion of ^{111}indium labeled leucocytes[1] and with both microscopic and endoscopic evidence of GI inflammation[2,3]. In addition to its use in differentiating irritable bowel syndrome (IBS) from inflammatory bowel disease (IBD)[4-8], it has been assessed as a marker of mucosal healing[9-12]. There are seven published studies in adult IBD patients that address the issue of FC as a predictor of clinical relapse with ongoing medical therapy in quiescent Crohn's disease (CD)[13-19]. Higher FC levels were associated with a greater risk of relapse for those with Ulcerative colitis (UC)[13-17], but discrepant results have been seen in CD[14,15,19]. Furthermore meta-analysis has shown that there is insufficient evidence to determine whether FC levels in those with ileal CD can serve to predict relapse[20]. The aim of this study was to prospectively assess the role of FC as a predictive marker of relapse within 12 months in those with asymptomatic CD of all phenotypes.

METHODS

Patients

In this single centre prospective study, 97 consecutive CD patients in clinical remission attending for routine outpatient review between August 2010 and November 2011 were

identified and enrolled. Written informed consent was obtained at time of enrollment. Remission was defined as a Crohn's disease Activity Index (CDAI)[21] of <150 points.

We excluded patients with an unclear diagnosis (ie. 'indeterminate colitis'), clinical relapse within the preceding 3 months, concomitant serious illness, pregnancy, age<18 years, alcohol abuse, non steroidal anti-inflammatory use, and stool culture positivity.

Full ethical approval was awarded on 15 April 2010 by the West of Scotland Research Ethics Service (WeSRES) (REC reference 10/S0704/1).

The first FC sample provided by each participant from our recently published study[22] on the intra-individual variability of FC was used as a baseline value for this prospective follow up study. The samples were collected by the patients at home and processed at the biochemistry laboratory at Glasgow Royal Infirmary. Patients were reviewed at regular 3-6 monthly intervals or at relapse. The primary endpoint was relapse within 12 months, while the secondary endpoint was relapse at any time during follow-up. Relapse was defined as an unplanned escalation in therapy, progression of disease phenotype by the Montreal classification, or hospitalization and/or emergency surgery for active CD.

Biochemistry procedures

The Roche faecal extraction device was used to prepare and analyze stool samples adhering to the manufacturers instructions (Bühlmann calprotectin ELIZA kit). Stool was collected in screw-capped plastic containers and received by laboratory within 48 hours of the final stool collection. Samples were processed by qualified biochemical scientists with Health professionals council registration on site at Glasgow Royal Infirmary. Samples weighing.

between 98–102 mg were placed into the extraction tube cap, 4.9ml of extraction buffer was then added to all tubes which were recapped and homogenised for 15 minutes on the Alpha multi tube vortexer at maximum speed. The homogenate was centrifuged at 3000rpm for 10 minutes and the supernatants were transferred to plastic tubes and stored at -20^{o}C. The time from sampling to preparation and freezing was approximately 1 to 3 days. The supernatants were thawed then mixed and centrifuged before analysis with the Bühlmann quantitative calprotectin ELISA kit on the Triturus automated ELISA analyser for determination of calprotectin concentration in stools. Calprotectin was expressed as micrograms per gram (µg/g) of faeces. The faecal samples were stable between 2-8^{o}C for up to 10 days and faecal extracts for 4 months at -20^{o}C.

Statistical considerations

The Mann-Whitney or t-test was used, as appropriate, to test for significant differences in continuous variables (including FC) between patients who relapsed by 12 months and those who did not, while Fisher's exact test was used for categorical variables.

The sensitivity and specificity of different FC values to predict relapse by 12 months were calculated for all those who either reached the primary endpoint within 12 months or were followed up for at least 12 months without reaching the primary endpoint, and the resulting receiver operating characteristic (ROC) curve plotted. The corresponding area under the curve (AUC) was calculated to represent the overall predictive power of FC in predicting relapse up to 12 months later. The sensitivity, specificity, and negative and positive predictive values are presented for the FC cutoff value with the optimal balance of sensitivity

and specificity. Patients who died or were otherwise lost to follow-up before relapsing or being followed up for 12 months were excluded from this part of the analysis.

The optimal FC cutoff value was subsequently used to calculate Kaplan-Meier (K-M) cumulative event curves of time to relapse for all patients throughout the entire study. Patients who did not relapse were censored at end of follow-up. A Cox proportional hazards model was fitted to assess the impact of a FC value above or below the chosen cutoff on time to relapse at any point in the study, adjusted for age (in years), gender, any previous surgery (yes/no), and stoma (yes/no).

Sample size calculation

The sample size was calculated for the reliability phase of this study, which is reported in detail elsewhere[22]. Briefly, we estimated that 95 patients would have 80% power to show a 95% confidence interval of total width 0.13 around an intraclass correlation coefficient of 0.9 between the FC values from 3 samples.

RESULTS

The mean age of all 97 recruited patients at baseline was 47 years (SD 16), 38% were male and 20% were smokers. Montreal Classification of CD was as follows: age at diagnosis (A1 8%, A2 71%, A3 21%), location (L1 16%, L2 36%, L3 47%) and behavior (B1 59%, B2 30%, B3 11%, p 15%).

Of 97 patients recruited, the care of three individuals was transferred to another centre, one died of non-IBD related pathology without reaching an endpoint and one was lost to follow

up prior to reaching either the primary endpoint or a follow up of 12 months. The sensitivity/specificity part of the analysis therefore included 92 patients.

Of these 92 patients, 10 (10.9%) relapsed within 12 months. Table 1 shows that patients who experienced a relapse within 12 months exhibited higher median FC levels at baseline (414µg/g; IQR 259-590) than those who did not (96 µg/g; IQR 39-237; p=0.005). There were no significant differences in age, gender, surgery, stoma, smoking, age at diagnosis, location, behavior or CRP between those who did and did not relapse by 12 months. Patients were more likely to relapse if they were taking chronic steroids, but very small numbers were on steroids and the difference was only marginally significant. There were no other relationships between drugs and relapse by 12 months.

Figure 1 shows the ROC curve for predicting relapse by 12 months, with sensitivity and specificity of various cutoff values of FC. The optimal balance of sensitivity and specificity corresponded to an FC cutoff value of 240µg/g. This cutoff gave a sensitivity of 80.0%, specificity of 74.4%, negative predictive value of 96.8% and positive predictive value of 27.6%. The area under the curve (AUC) to predict CD relapse at 12 months using FC determination was 77.4%.

Figure 2 explains the discrepancy between the negative and positive predictive values. Very few patients who relapsed by 12 months had FC of less than 240µg/g. Therefore if a patient had FC<240µg/g they would have a low risk of relapse over the ensuing 12 months. However, while many more patients with FC≥240µg/g did relapse, two-thirds of all patients with FC of or above 240µg/g did not and this is reflected in the low positive predictive value for relapse prediction in our cohort.

The selected cutoff of 240µg/g was used to produce Kaplan-Meier (K-M) cumulative event curves of time to relapse for all 97 patients and these are presented in Figure 3. A total of 15 patients relapsed throughout the study. The shortest time to relapse was 87 days and the longest was 560 days (1.5 years), while the remaining 82 patients were followed up for between 98 (3 months) and 849 days (2.3 years) without relapsing. There is a clear separation between the curves, with patients with FC≥240µg/g having a substantially shorter time to relapse than those with FC below the cutoff.

Table 2 shows the results of the Cox proportional hazards model of FC on time to relapse adjusted for demographics. The model confirmed the difference in time to relapse between those with high or low FC exhibited in the K-M curves, with a hazard ratio (HR) for FC≥240µg/g vs FC<240µg/g of 5.7 (95% CI 1.9-17.3; p=0.004). Thus, based on our sample, a patient with FC≥240µg/g is around two to 17 times more likely to relapse within about two years than one with FC below 240µg/g. Table 2 also shows that there was no impact of demographics on time to relapse.

As exploratory analyses, the 16 patients with ileal only disease and the 35 patients with colonic only disease were considered separately. Three of the 16 ileal patients relapsed during the study, at 142, 394 and 560 days, with only one having relapsed by 12 months. The three ileal patients who relapsed at any time during the study had higher median baseline FC levels (371 µg/g; IQR 284-741) than the 13 who did not (57 µg/g; IQR 20-101). A Mann Whitney test showed marginal statistical non-significance for this difference (P=0.057). Two of the 35 colonic patients relapsed by 12 months, at 94 and 298 days, with a further two relapsing later at 470 and 524 days. The four colonic patients who relapsed at any time

during the study had higher median baseline FC levels (424 µg/g; IQR 209-695) than the 31 who did not (187 µg/g; IQR 48-386), though the difference was not statistically significant (p=0.16).

DISCUSSION

Our prospective dataset, which is the largest yet studied, demonstrates that an FC concentration below 240µg/g is predictive of a low risk of clinical relapse within 12 months for adults with quiescent CD. The ROC curve analysis revealed that the 240µg/g concentration gave an optimal balance of sensitivity (80.0%) and specificity (74.4%). Despite a reassuringly high negative predictive value of 96.8% at this cut off, the low positive predictive value (27.6%) suggested that FC is most useful as a tool to predict low risk of clinical relapse.

FC is a relatively cheap and non-invasive test making its use attractive in the increasingly financially conscious and risk-averse realm of modern health care. Its use is now established in differentiating irritable bowel syndrome (IBS) from inflammatory bowel disease (IBD)[4-8]. There is also data supporting its use in evaluating abdominal discomfort[23], reducing the need for endoscopy in suspected IBD[24], assessing treatment response in IBD[25,26], predicting mucosal healing in IBD[9,12,26-27], detecting post operative relapse in CD[28-30] and predicting response to anti-Tumour necrosis factor (anti-TNF) therapy[31-33.]

We have recently shown that FC levels do not vary considerably within individuals with quiescent CD on a day to day basis[22]. This is reassuring if a one off FC is to be used as a tool to predict future relapse. It would be beneficial for clinicians to target early effective therapies if they could better predict risk. It has been noted that FC may correlate more

closely with endoscopic scores than clinical severity scores in CD[34], hinting at the potential to predict preclinical disease. There are published data on the use of FC levels to predict relapse in UC, which consistently show it to be both sensitive and specific[13-17, 20].

There are seven previously published prospective studies that have explored the issue of utilising FC concentrations to predict relapse within a 12 month time period in quiescent adult CD, showing conflicting results[13-19]. Three studies showed no statistically significant difference between the baseline median FC for relapsers and non-relapsers[14,15,19]. Costa et al[14] compared FC levels in both UC patients and 38 CD patients[14]. Although the results were higher for those relapsing in UC, the levels in CD were comparable for both relapsers and non-relapsers (220.1 vs 220.5µg/g P=0.395). Similarly, in the 65 CD patients studied by D'Inca et al[15] there was also no statistically significant difference between relapse and non-relapse median FC levels (207 vs 88 mg/kg P=0.55). They found that the subgroup of colonic CD were the only group where FC level was predictive, but the numbers of relapsers in this cohort were small (4 of 6 colonic patients relapsed). The study by Laharie et al[19] differed from the others described, as their 50 CD patients were all in remission 14 weeks post infliximab induction. They found no significant difference between week 14 FC levels in relapsers vs non-relapsers (200 vs 150µg/g P=NS). This cohort had a high 12-month relapse rate (46% vs 12% in our study).

Conversely, four other studies[13,16-18] showed a positive association between FC baseline level and risk of relapse. Tibble and colleagues[13] use a different and older assay but the FC results are equivalent to those of later studies (calprest) when the result is multiplied by a factor of five. They showed in 43 CD patients that relapsing patients had higher median baseline FC compared with that of non-relapsers (122mg/l – which converts to 610µg/g vs 42mg/l – which converts to 220µg/g). They combined the results with UC patients to produce a

receiver operator curve (ROC) curve showing that an equivalent FC concentration of over 250µg/g predicted relapse with a sensitivity of 90% and a specificity of 83%. Garcia-Sanchez et al[17] studied 66 CD patients and identified a best cut off value of 200µg/g (sensitivity 80%, specificity 65%, PPV 46%, and NPV 88%) to predict relapse. They did, however stress that relapse predictability was more accurate for colonic CD. Their PPV and NPV values were similar to our own which suggest greater accuracy for prediction of remission than relapse. In the study by Gisbert et al[16], a total of 89 CD patients were included, 13 of whom relapsed. FC levels were found to be higher in relapsers (266 vs 145µg/g; P=0.002). Both Gisbert[16] and Garcia-Sanchez[17] drew attention to the fact that a high FC level appeared to be more predictive of relapse in colonic disease. Published commentaries[35-37] suggested that two of the earlier studies[13,14] had conflicting results which could be accounted for by differing proportions of small bowel CD patients. Given that greater levels of excreted indium^{111}-labelled leucocytes have been found in colonic vs small bowel CD[38], Kallel et al[13] were prompted to exclude small bowel CD from their analysis of 53 CD patients. Higher median FC values were measured at baseline in the relapse group (380.5µg/g cf 155µg/g P<0.001). The ROC curve analysis revealed that a level of >340µg/g provided the maximal sum of sensitivity (80%) and specificity (90.7%) to predict relapse.

More recently, Primas and colleagues[39] have published an abstract describing 57 CD patients post ileocolonic resection. They found that an FC cut off of over 100µg/g 6 months post surgery could predict relapse at a median 11 months post surgery with a sensitivity of 93% and specificity of 47%. Additionally, a large retrospective analysis of 650 patients (32% relapsers) by Kennedy et al[40] has been published in abstract form; the primary endpoint was a composite of Montreal behaviour progression, hospitalisation for a flare of disease or surgery. A total of 211 reached the endpoint within 12 months of whom 57 had a progression in Montreal behaviour. They discovered a significant difference between median FC levels in

relapsers vs non-relapers (595 vs 320µg/g). It should be noted that neither Primas nor Kennedy appear to have identified clinically quiescent patients at baseline as was done in the seven published studies described[13-19]. The patient populations may therefore not be directly comparable.

Louis et al[41] studied a very different CD population who had undergone at least 12 months infliximab therapy in combination with an antimetabolite. Anti-TNF therapy was withdrawn and calprotectin was measured at 2 monthly intervals. Of the 115 patients studied, 85 had FC measurements. An FC level >300µg/g at baseline was associated with relapse (hazard ratio estimate 2.5 p=0.04). Desuray et al[42] looked at this data in greater detail in 113 patients, finding a sharp rise in FC within 4 months of relapse with a FC cut off of 305µg/g giving a sensitivity of 70% and specificity of 74% for relapse prediction.

Our own study shows a positive association between FC level and risk of relapse. Our relapse rate is relatively low at 11% compared with 18.9% in the group studied by Kallel et al[19] and 58% in the study by Tibble et al[13]. Our study, and that by Gisbert et al[16], included patients on continuing biological therapy which may contribute to their similarly low relapse rate (14.6%). We chose to study all CD phenotypes in an attempt to establish, in a larger cohort, whether an effective cutoff could be determined. Our most effective FC cut off level to predict relapse was shown to be 240µg/g, which is similar to that of Tibble et al[13] but lower than the 340µg/g of Kallel et al[19]. Given that the latter study[19] excluded those without colonic disease and indium^{111} leucocytes are excreted in higher levels in colonic disease[38], this could explain the higher cut off level.

In the meta-analysis by Mao et al[20], it was commented that there were insufficient available data to determine the use of FC to predict relapse in ileal CD. It is more challenging to assess proximal gastro intestinal inflammation by endoscopy and patients with inflammatory ileal

disease may be less likely to have symptoms than those with colonic CD, leading to a greater chance of progressing to fistulising or stricturing disease[43]. Thus, although FC levels tend to be lower in ileal as opposed to colonic disease[38], the FC test may well have a greater discriminant ability in ileal CD due to the disconnect between clinical symptoms and ileal disease activity in this cohort. Interestingly, we found that the difference in median baseline FC between relapsers and non-relapsers with ileal CD was large and close to statistical significance (371 vs 57µg/g; p=0.057) despite the small number of patients in this subgroup. It should be noted that these figures were obtained over an extended period of about two years, during which time 3 of the 16 ileal patients had relapsed, since only 1 patient relapsed within 12 months. This result suggests that a larger study, perhaps over a longer time period, of those with ileal CD would be worthwhile to clarify this potential association. In our subgroup of colonic CD, only two of 31 patients relapsed within 12 months and four in the whole study period. Although the mean FC values were higher for the four relapsers (424 vs 187µg/g), statistical significance was not reached (p=0.16).

Although some published articles have shown an association between higher baseline C reactive Protein (CRP) and risk of relapse[18,42,43], others have failed to show this[13,14,16]. This study also shows no association between baseline CRP and the risk of subsequent clinical relapse. It should, however, be noted that this study was not designed to detect such an association and approximately half of the patients did not have a baseline CRP measurement. We did not show an association with relapse and smoking, but the numbers of those smoking at baseline were low (16) and we did not specifically collect data on starting or stopping smoking during the study.

There are additional limitations in our study which merit consideration. Like the previous studies described, we used CDAI rather than endoscopy as an objective assessment of disease

activity to define remission. This measure has been shown to correlate poorly with more objective assesments[12]. We would argue however that the risk of endoscopy is not justified for those in remission as it does not reflect current clinical practice. Furthermore, we conducted the study in a tertiary referral institute that may make the findings less applicable to the general population, although no novel therapies are used. The investigators were not blinded to the FC results, but investigators only based treatment decisions on their clinical assessment as there is no prior evidence to change treatment due to FC alone.

In conclusion, our study, utilising the largest prospective dataset in the current literature, provides clearer evidence that adults with quiescent CD with a faecal calprotectin level below 240µg/g are unlikely to relapse within 12 months, while those with a level of 240µg/g or above are substantially more likely to relapse within 12 months. This information can be obtained by non-invasive means and can provide both prognostic information for patient and clinician and a therapeutic target for physicians treating Crohn's patients who are in clinical remission when attending the outpatient clinic.

STATEMENT OF INTEREST

The biochemistry processing and statistical analysis required financial support which was provided by an educational grant from 'Cure Crohn's colitis' (C^3) www.curecrohnscolitis.org - charity number SC036559

ACKNOWLEDGEMENTS

Dr Naismith and Dr Gaya designed the study. Dr Rankin coordinated processing of samples. Ms Munro, Ms Laird, Dr Morris, Dr Winter, Dr Gaya and Dr Smith collected the data. Dr Smith collated the data. Dr Barry performed statistical analysis and wrote the results and statistical methods sections of the paper. Dr Naismith wrote the paper with contributions from Dr Smith, Dr Gaya and Dr Rankin.

STUDY HIGHLIGHTS

What is Current Knowledge

Faecal calprotectin is useful to predict disease recurrence in ulcerative colitis

Faecal calprotectin has shown conflicting results when utilised as a predictive marker in quiescent Crohn's

Faecal calprotectin is thought to be less useful to predict relapse in ileal Crohn's disease

What is new here?

Confirmation of the predictive value of faecal calprotectin in Crohn's in the largest prospective study

Faecal Calprotectin shows potential to predict disease recurrence in quiescent ileal Crohn's

REFERENCES

1. Roseth AG, Schmidt PN, Fagerhol MK. Correlation between faecal excretion of indium-111-labelled granulocytes and calprotectin, a granulocyte marker protein in patients with inflammatory bowel disease. *Scand J Gastroenterol* 1999;**34**:50–4.

2. Roseth AG, Aadland E, Jahnsen J et al. Assessment of disease activity in Ulcerative Colitis by faecal calprotectin, a novel granulocyte marker protein. *Digestion* 1997;58:176-180

3. Sipponen T, Saviahti E Kohlo KL et al. Crohn's disease activity assessed by fecal calprotectin and lactoferrin: correlation with Crohn's disease activity index and endoscopic findings. *Inflam Bowel Dis*. 2008;**14**:40-46

4. Tibble JA, Teahon K, Thjodleifsson B, *et al*. A simple method for assessing intestinal inflammation in Crohn's disease. *Gut 2*000;**47**:506–13.

5. Sutherland AD, Geary RB, Frizelle FA. Review of fecal biomarkers in inflammatory bowel disease. *Dis Colon Rectum* 2008; **51**:1283-91

6. Schoepfer AM, Trummler M, Seeholzer P, Seibold-Schmid B, Seibold F. Discriminating IBD from IBS: comparison of the test performance of fecal markers, blood leucocytes, CRP and IBD antibodies. *Inflamm Bowel Dis* 2008;**14:**32-9

7. Gisbert JP, McNicholl AG. Questions and answers on the role of faecal calprotectin as a biological marker in inflammatory bowel disease. *Dig Liver Dis* 2009;**41**:55-66

8. Centre for Evidence-based purchasing. Value of Calprotectin in screening out Irritable Bowel Syndrome: Evidence Review. London: NHS Purchasing and Supply Agency, 2009: CEP09026

9. Roseth AG, Aadland E, Grzyb K. Normalization of faecal calprotectin: a predictor of mucosal healing in patients with inflammatory bowel disease. *Scand J Gastroenterol* 2004;**39:**1017–20.

10. Langhorst J, Elsenbruch S, Koelzer J, Rueffer A, Michalsen A, Dobos G J. Noninvasive markers in the assessment of intestinal inflammation in inflammatory bowel disease: performance of fecal lactoferrin, calprotectin and PMN-elastase, CRP and clinical indices. *Am J Gastroenterol* 2008;**103**:162-169

11. Jones J, Loftus EV, Panaccione R et al. Relationships between disease activity and serum and fecal biomarkers in patients with Crohn's disease. *Clin Gastroenterol Hepatol* 2008;**6**:1218-24

12. Schoepfer AM, Beglinger C, Straumann A et al. Fecal Calprotectin correlates more closely with with simple endoscopic score for Crohn's disease (SES-CD) than CRP, Leukocytes and the CDAI. *Inflamm Bowel Dis* 2010;**105**:162-169

13. Tibble J.A, Sigthorsson G., Bridger S., Fagerhol M.K., Bjarnason I. Surrogate Markers of Intestinal Inflammation Are Predictive of Relapse in Patients With Inflammatory Bowel Disease. *Gastroenterology* 2000; **119**: 15-22.

14. Costa F, Mumolo MG, Ceccarelli L*et al*. Calprotectin is a stronger predictive marker of relapse in ulcerative colitis than in Crohn's disease. *Gut* 2005;**54**:364-368

15. D'Inca R, Dal Pont E, Di Leo V *et al*. Can Calprotectin predict relapse risk in inflammatory bowel disease? *Am J Gastroenterol* 2008;**103**:2007-2014

16. Gisbert JP, Bermejo F, Perez-Calle JL *et al*. Fecal calprotectin and lactoferrin for the prediction of inflammatory bowel disease relapse. *Inflamm Bowel Dis.* 2009;**15**(8):1190–8.

17. Garcia-Sanchez V, Iglesias-Flores E, Gonzalez R *et al*. Does Fecal calprotectin predict relapse in patients with Crohn's disease and ulcerative colitis? *J Crohns Colitis* 2010 **4**(2):144-52

18. Kallel L, Ayadi I, Matri S *et al*. Fecal calprotectin is a predictive marker of replase in Crohn's disease involving the colon: a prospective study. *European Journal of Gastroenterology & Hepatology* 2010;**22**:340-345

19. Laharie D, Mesli S, El Hajbi F *et al*. Prediction of Crohn's disease relapse with faecal calprotectin in infliximab responders: a prospective study. *Aliment Pharmacol Ther* . 2011;**34**(4):462-9

20. Mao R, Xiao YL, Gao X *et al*. Fecal calprotectin in predicting relapse of inflammatory bowel diseases: a meta-analysis of prospective studies. *Inflamm. Bowel Dis.*2012;**18**(10),1894–1899

21. Best WR, Becktel JM, Singleton JW, Kern F Jr *et al*. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. *Gastroenterology*. 1976; **70**: 439–444

22. Naismith GD, Smith LA, Barry SJE, Munro JI, Laird S, Rankin K, Morris AJ, Winter JW, Gaya DRG. A prospective single centre evaluation of the intra-individual variability of the faecal calprotectin test in quiescent Crohn's disease. *Aliment Pharmacol Ther* 2013;**37**:613-21

23. Manz M, Burri E, Rothen C. *et al*. Value of fecal calprotectin in the evaluation of patients with abdominal discomfort: an observational study. *BMC Gastroenterol*. 2012;**12**:5

24. Van Rheenen PF, Van de Vijver E, Fidler V. Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. *BMJ* 2010;**341**:c3369.

Formatted: Italian (Italy)

25. Wagner M, Peterson GB, Ridefelt P, Sangfelt P, Carlson M. Fecal markers of inflammation used as surrogate markers for treatment outcome in relapsing inflammatory bowel disease. *World J Gastroenterol* 2008;**14**(36)5584-5589

26. Langhorst J, Elsenbruch S, Koelzer J, Rueffer A, Michalsen A, Dobos G J. Noninvasive markers in the assessment of intestinal inflammation in inflammatory bowel disease: performance of fecal lactoferrin, calprotectin and PMN-elastase, CRP and clinical indices. *Am J Gastroenterol* 2008;**103**:162-169

27. Jones J, Loftus EV, Panaccione R et al. Relationships between disease activity and serum and fecal biomarkers in patients with Crohn's disease. *Clin Gastroenterol Hepatol* 2008;**6**:1218-24

28. Lamb CA, Mohiuddin MK, Gicquel J, et al. Faecal calprotectin or lactoferrin can identify postoperative recurrence in Crohn's disease. *Br J Surg* 2009; **96**: 663–74.

29. Boschetti, G., D. Moussata, A-L.Charlois, B. Flourié, and S. Nancey. P314 Accuracy of fecal calprotectin in the assessment of post-operative endoscopic recurrence in patients with Crohn's disease. *Journal of Crohn's and Colitis* 7 (2013): S134-S135.

30. Beltrán, B., E. Cerrillo, M. Iborra, I. Moret, F. Rausell, L. Tortosa, G. Bastida, M. Aguas, and P. Nos. "P172 Fecal calprotectin (FC) is a useful early predictive marker for postoperative recurrence in Crohn's disease (CD)." *Journal of Crohn's and Colitis* 6 (2012): S78

31. Guidi, L, Marzo M, Andrisani G, Felice C, Pugliese D, De Vitis I, Papa A, Rapaccini GL, Forni F, Armuzzi A. P214 Faecal calprotectin assay after induction with anti-TNF alpha agents in IBD patients: prediction of clinical response and mucosal healing at one year. *Journal of Crohn's and Colitis* 7 (2013): S95.

32. Ferreiro, R., M. Barreiro-de Acosta, M. Otero, A. Lorenzo, C. Alonso, and J. E. Dominguez-Munoz. P109 Use of faecal calprotectin as predictor of relapse in patients under maintenance treatment with infliximab. *Journal of Crohn's and Colitis* 7 (2013): S52-S53.

33. Molander, P, af Björkesten C. G, Mustonen, H, Haapamäki J, Vauhkonen M, Kolho et al. (2012). Fecal calprotectin concentration predicts outcome in inflammatory bowel disease after induction therapy with TNF± blocking agents. *Inflammatory bowel diseases*, *18*(11), 2011-2017

34. **D'Haens G, Ferrante M, Vermeire S, Baert F, Norman M, Moortgat L et al. Fecal calprotectin is a surrogate marker for endoscopic lesions in inflammatory bowel disease.** *Inflammatory bowel diseases*, 18(12):2218-24, 2012 Dec.

35. Hanaway P, Roseth A Inflammatory biomarkers predict relapse in IBD. *Gut* 2005;**54**:1346-1347

Formatted: English (United Kingdom)

36. Pardi D S. Sandborn, W. J. Predicting relapse in patients with inflammatory bowel disease: what is the role of biomarkers?. *Gut* 2005 **54**(3), 321-322.

37. Lemann M, Mary JY. Calprotectin and IBD. *Gut* 2005; **54**:1349-1350

38. Saverymuttu SH, Peters AM, Crofton ME 111 Indium autologous granulocytes in the detection of inflammatory bowel disease. Gut 1985; **26**:955-960

39. Primas, C, Frühwald G, Angelberger S, Allerstorfer D, Papay P, Eser A, Gratzer C et al. P381 Role of fecal calprotection in predicting ileocolonic endoscopic recurrence in postoperative Crohn's disease. *Journal of Crohn's and Colitis* 7 (2013): S162.

40. Kennedy, N. A, Chang J, Guy MH, Smith T, Loh JT, Haunschmidt D, Muscat M et al. P250 Elevated faecal calprotectin predicts disease progression in Crohn's disease. *Journal of Crohn's and Colitis* 7 (2013): S109-S110.

41. Louis, E, Mary, J Y, Vernier–Massouille, G, Grimaud J C, Bouhnik Y, Laharie D et al (2012). Maintenance of remission among patients with Crohn's disease on antimetabolite therapy after infliximab therapy is stopped. *Gastroenterology*, *142*(1), 63-70

42. De Suray, Nicolas, Julia Salleron, Gwenola Vernier-Massouille, Jean-Charles Grimaud, YoramBouhnik, David Laharie, Jean-Louis Dupas et al. "864 Close Monitoring of CRP and Fecal Calprotectin is Able to Predict Clinical Relapse in Patients With Crohn's Disease in Remission After Infliximab Withdrawal. a Sub-Analysis of the Stori Study." *Gastroenterology* 142, no. 5 (2012): S-149.

43. Cosnes, J., Cattan, S., Blain, A., Beaugerie, L., Carbonnel, F., Parc, R. and Gendre, J.-P. (2002), Long-term evolution of disease behavior of Crohn's disease. *Inflamm Bowel Dis*, **8**: 244–250

STARD CHECKLIST

Section and Topic	Item #		On page #
TITLE/ABSTRACT/ KEYWORDS	1	Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	2,4
INTRODUCTION	2	State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	2,4
METHODS			
Participants	3	The study population: The inclusion and exclusion criteria, setting and locations where data were collected.	4,5
	4	Participant recruitment: Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	4
	5	Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	Yes,4
	6	Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	Before - prospective
Test methods	7	The reference standard and its rationale.	4,5
	8	Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	4,5
	9	Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	2,4,6
	10	The number, training and expertise of the persons executing and reading the index tests and the reference standard.	5
	11	Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other test and describe any other clinical information available to the readers.	Not blinded
Statistical methods	12	Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g. 95% confidence intervals).	5,6
	13	Methods for calculating test reproducibility, if done.	5,6
RESULTS			
Participants	14	When study was performed, including beginning and end dates of recruitment.	4
	15	Clinical and demographic characteristics of the study population (at least information on age, gender, spectrum of presenting symptoms).	Table 1
	16	The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to undergo either test (a flow diagram is strongly recommended).	7
Test results	17	Time-interval between the index tests and the reference standard, and any treatment administered in between.	4
	18	Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	Table1
	19	A cross tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	Fig1-3
	20	Any adverse events from performing the index tests or the reference standard.	None n/a
Estimates	21	Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g. 95% confidence intervals).	5,6,7,8
	22	How indeterminate results, missing data and outliers of the index tests were handled.	5,6,7,8
	23	Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centers, if done.	5-9
	24	Estimates of test reproducibility, if done.	
DISCUSSION	25	Discuss the clinical applicability of the study findings.	9-14

FIGURE LEGENDS

Figure 1 Receiver operating characteristic (ROC) curve of the sensitivity and specificity of FC predicting relapse at 12 months for the 93 patients followed up for at least that length of time, based on various cutoffs of FC

Figure 2 Scatterplot of the FC values of all patients in the study, with those who relapsed by 12 months marked in red, and those who did not marked in black. The optimal cutoff of 240μg/g marked as a dashed line

Figure 3 Kaplan-Meier (K-M) cumulative event curves of time to relapse in days for all patients in the study, stratified by whether their FC was below or above 240μg/g. Patients who did not relapse were censored at end of follow-up and are marked by crosses on the K-M curves